

Does hospital overload increase the risk of death when infected by SARS-CoV-2?

Benjamin Glemain^{1,2*}, Charles Assaad¹, Walid Ghosn³,
Paul Moulaire¹, Xavier de Lamballerie⁴, Marie Zins^{5,6},
Gianluca Severi^{7,8}, Mathilde Touvier⁹, Jean-François Deleuze¹⁰,
SAPRIS-SERO study group, Nathanaël Lapidus^{1,2†},
Fabrice Carrat^{1,2†}

¹Sorbonne Université, Inserm, Institut Pierre-Louis d'épidémiologie et de santé publique, Paris, France.

²Département de santé publique, Hôpital Saint-Antoine, APHP, Paris, France.

³Centre d'épidémiologie sur les causes médicales de décès de l'Inserm, Inserm-CépiDc, Paris, France.

⁴Unité des Virus Émergents, UVE, Aix Marseille Univ, IRD 190, INSERM 1207, IHU Méditerranée Infection, Marseille, France.

⁵Paris University, Paris, France.

⁶Université Paris-Saclay, Université de Paris, UVSQ, Inserm UMS 11, Villejuif, France.

⁷CESP UMR1018, Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, Villejuif, France.

⁸Department of Statistics, Computer Science and Applications, University of Florence, Italy.

⁹Sorbonne Paris Nord University, Inserm U1153, Inrae U1125, Cnam, Nutritional Epidemiology Research Team (EREN), Epidemiology and Statistics Research Center – University of Paris (CRESS), Bobigny, France.

¹⁰Fondation Jean Dausset-CEPH (Centre d'Etude du Polymorphisme Humain), CEPH-Biobank, Paris, France.

*Corresponding author(s). E-mail(s): benjamin.glemain@inserm.fr;

†These authors contributed equally to this work.

Abstract

Several studies found an association between the risk of death for COVID-19 patients and hospital overload during the first pandemic wave. We studied this association across the French departments using 82,467 serological samples and a hierarchical Bayesian model. In high-incidence areas, we hypothesized that hospital overload would increase infection fatality rate (IFR) without increasing infection hospitalization rate (IHR). We found that increasing departmental incidence from 3% to 9% rose IFR from 0.42% to 1.14%, and IHR from 1.66% to 3.61%. An increase in incidence from 6% to 12% in people under 60 was associated with an increase in the proportion of people over 60 among those infected, from 11.6% to 17.4%. Higher incidence did increase the risk of death for infected persons, probably due to an older infected population in high-incidence areas rather than hospital overload.

Keywords: COVID-19, Hospital overload, Bayesian statistics, Hierarchical modeling, Spatial modeling, Causal graph

1 Introduction

1 The first wave of the COVID-19 pandemic led to episodes of hospital overload in many
2 parts of the world, requiring the urgent provision of additional hospital beds to pre-
3 vent excess mortality [1, 2]. Temporary units were set up, but suspected of being less
4 effective than permanent units (although these comparisons were impeded by con-
5 founders such as admission criteria) [3–5]. Indeed, hospital overload can manifest in
6 several ways. First, it can appear as a lack of hospital bed availability, resulting in a
7 lower proportion of infected individuals being hospitalized. Second, it can lead to a
8 decrease in the quality of care due to the urgent addition of extra beds with poten-
9 tially inadequate facilities, equipment, or staff. Hospital overload thus has multiple
10 aspects, making it difficult to measure using a single indicator.
11 So far, the consequences of these episodes of hospital overload on the risk of death for
12 infected persons have mainly been studied through the number of confirmed COVID-
13 19 cases, using the case fatality rate (CFR) [6–8]. Hospital overload was notably
14 expressed as the ratio between the number of COVID-19 cases and baseline hospital
15 resources, a measure that does not distinguish between situations of high incidence
16 and those of insufficient hospital resources at baseline [6]. In addition, CFR analysis is
17 limited by spatial and temporal variations in the case detection rate (the proportion
18 of new cases that are detected) [9, 10].
19 To overcome the limitations of the cases-based approach, several studies used serolog-
20 ical data and estimated infection fatality rates (IFRs) [11–14]. IFR is defined as the
21 number of deaths attributed to COVID-19 during a given period divided by the num-
22 ber of infected individuals (incidence) over the same period, and corresponds to the
23 risk of death for infected persons. These studies compared IFRs across countries and
24 age groups, but did not explore the impact of hospital overload.
25 Built on the seroprevalence study SAPRIS-SERO, the present work aimed to estimate
26

27 the effect of COVID-19 incidence on IFR at the scale of the metropolitan depart-
28 ments of France during the first pandemic wave in the population over 20. We used a
29 Bayesian statistical framework to leverage multiple sources of data and to account for
30 uncertainty surrounding the latent variables when used in regressions (such as inci-
31 dence and IFR). As incidence (understood as a cumulative incidence over the first
32 wave) was reported to range from 3% to 9% in the French regions (administrative sub-
33 divisions gathering several departments), the consequences of an incidence shift from
34 3% to 9% on IFR were the main object of this study [15, 16].

35 The relation between incidence and IFR is however confounded, notably because the
36 determinants of IFR may share socio-economic causes with incidence at the scale of
37 departments. Typically, wealthier departments could have a population which travels
38 more (possibly increasing incidence) and which is healthier (decreasing IFR), partici-
39 pating in a spurious negative association between incidence and IFR. Thus, IFR was
40 adjusted (by conditioning and averaging, as described in the Methods section) for the
41 main determinants of COVID-19 outcome: prevalence of diabetes (as a surrogate for
42 obesity), proportion of the population over 60, and number of intensive care beds per
43 inhabitant (see the dedicated subsection: Choice of the covariates) [6, 7, 17, 18]. In
44 the Methods section (equation 15), we formalize the equivalence between (i) the asso-
45 ciation between incidence and adjusted IFR and (ii) the causal effect of incidence on
46 IFR, using the structural causal models framework.

47 Finally, the mechanisms by which incidence could influence IFR were explored. Evi-
48 dence for hospital overload with shortage of hospital beds was sought by examining
49 the effect of COVID-19 incidence on infection hospitalization rate (IHR), which is the
50 proportion of infected individuals being hospitalized for COVID-19. The age of the
51 infected individuals may also play a role in the effect of incidence on IFR. Indeed, it
52 has been suggested that older persons could be under-represented in the population
53 of infected people at the early stages of the epidemic [19]. To determine if infected
54 individuals were actually younger in departments with low incidence, the dependence
55 between COVID-19 incidence in people under 60 and the proportion of individuals
56 over 60 among those infected was investigated.

57 2 Results

58 2.1 Participants

59 The study included 82,467 persons with a serological test, living in metropolitan
60 France, and over 20 years old. All serological samples were collected between May and
61 November 2020. Among the participants, 319 reported a positive RT-PCR. These lat-
62 ter had a mean age of 52 years, 29% of them were males, and the median time elapsed
63 between RT-PCR and dried blood sampling was 111 days (IQR: 68-128). The partici-
64 pants with a positive RT-PCR were considered infected. On the other hand, 82,148
65 participants reported no positive RT-PCR (no RT-PCR or a negative RT-PCR). These
66 participants had a mean age of 58 years, and 35% of them were males. A detailed
67 flowchart and a table with the number of samples for each department and age group
68 are provided in the Supplementary information file.

69 2.2 Departmental incidence, IFR, and IHR

70 The model estimated an incidence of 6.8% over the first wave in metropolitan France, with a 95% credible interval (95% CI) of 6.4 to 7.2%. Figure 1 displays a map of

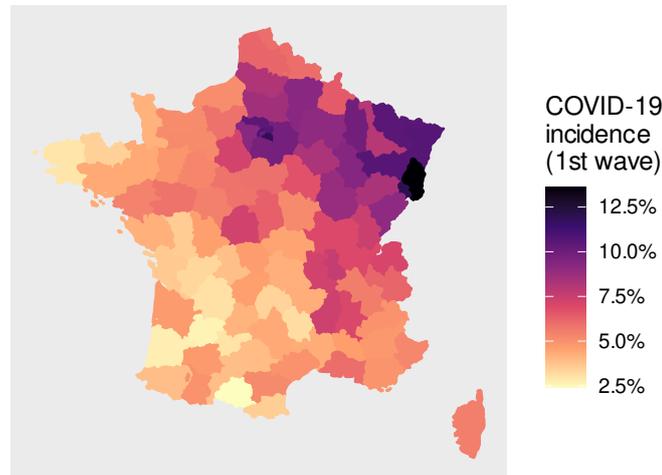


Fig. 1: COVID-19 departmental incidence (cumulated over the first wave) in metropolitan France.

71 departmental incidence, showing that north-east France was the most affected area.
72 Indeed, departmental incidence ranged from 2.5% (95% CI: 1.8-3.4%), in Ariège (a
73 south-west department), to 13.6% (95% CI: 12.1-15.2%), in Haut-Rhin (a north-east
74 department).

75
76 The overall infection fatality rate in metropolitan France was 0.94% (95% CI: 0.89-
77 0.99%) and infection hospitalization rate was 3.28% (95% CI: 3.10-3.46%). IFR ranged
78 from 0.27% (95% CI: 0.21-0.33%), in Haute-Garonne (a south-west department with
79 low COVID-19 incidence), to 1.97% (95% CI: 1.74-2.22%), in Haut-Rhin (the depart-
80 ment with the highest COVID-19 incidence, located in the north-east of France). IHR
81 ranged from 1.02% (95% CI: 0.72-1.42%), in Tarn-et-Garonne, to 5.92% (95% CI:
82 5.09-6.79%), in Territoire de Belfort. The Supplementary information file provides
83 exhaustive departmental estimates (incidence, IFR, IHR).

84 2.3 Effect of incidence on IFR and IHR

85 Figure 2 illustrates the association between departmental incidence and adjusted IFR,
86 or adjusted IHR (adjustment for the proportion of persons over 60 in the population,
87 for the prevalence of diabetes, and for the number of intensive care beds per inhab-
88 itant). An incidence of 3% was associated with an adjusted IFR of 0.42% (95% CI:
89 0.33-0.52%), and an incidence of 9% was associated with an adjusted IFR of 1.14%
90 (95% CI: 0.95-1.39%). The absolute difference (equivalent to the average causal effect

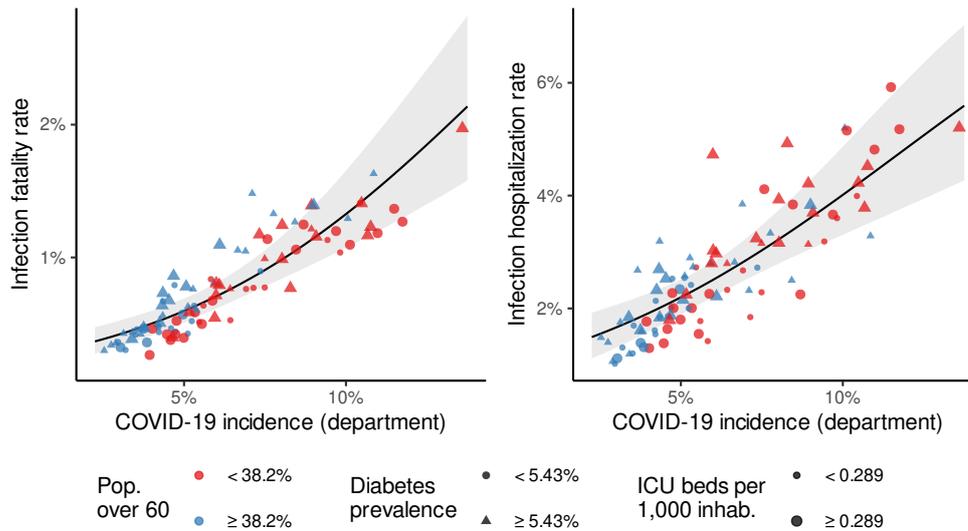


Fig. 2: Effect of departmental incidence on infection fatality rate (IFR) and on infection hospitalization rate (IHR).

The points represent the mean posterior departmental estimates.

Black line and gray zone: Posterior mean and 95% CI of the expected adjusted departmental IFR given incidence (or expected causal effect of incidence on IFR: see equation 15 of the Methods). The same description applies to IHR.

The covariates are represented relative to their medians.

Pop. over 60: Proportion of adult population over 60.

ICU beds per 1,000 inhab.: Number of intensive care beds per 1,000 inhabitants.

91 of an incidence shift from 3% to 9%) was 0.72% (95% CI: 0.49-1.01%).

92 An incidence of 3% was associated with an adjusted IHR of 1.66% (95% CI: 1.30-
93 2.06%), and an incidence of 9% was associated with an adjusted IHR of 3.61% (95% CI:
94 3.05-4.28%). The absolute difference was 1.94% (95% CI: 1.18-2.80%).

95 Complementary results (univariate analysis, role of the confounders) are provided in
96 the Supplementary information file.

97 2.4 Association between incidence and age of infected persons

98 As illustrated in Figure 3, a shift in incidence in the persons under 60 from 6% to 12%
99 (typical observed values) was associated with an increase in the expected proportion of
100 persons over 60 among those infected for a department with the same age structure as
101 metropolitan France, from 11.6% (95% CI: 9.6-13.6%) to 17.4% (95% CI: 15.5-19.5%).
102 The absolute difference was 5.8% (95% CI: 2.9-8.8%).

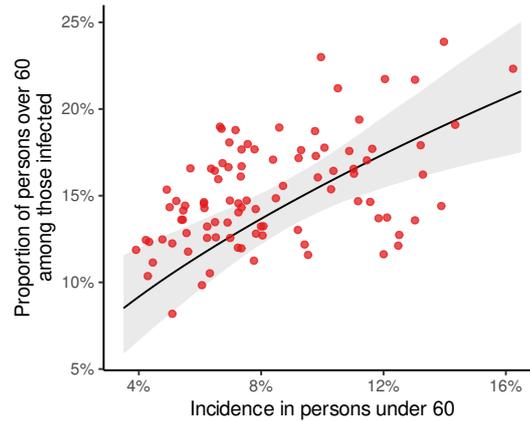


Fig. 3: Association between incidence in people under 60 and the proportion of people over 60 among those infected.

The points represent the mean posterior departmental estimates.

Black line and gray zone: Posterior mean and 95% CI of the expected proportion of persons over 60 among those infected (for a department with the same age structure as metropolitan France).

103 3 Discussion

104 This study explored the role of hospital overload on the risk of death for COVID-19
105 patients (IFR) using data collected in France following the first pandemic wave. We
106 found that a higher departmental incidence was associated with a higher adjusted
107 IFR, corresponding to a causal effect of incidence on IFR. This effect could possibly
108 be explained by the age of the infected persons, as we found a higher proportion of
109 people aged over 60 among those infected in high-incidence departments. The role of
110 hospital overload was explored through the analysis of the probability of hospitaliza-
111 tion when infected (IHR). In case of hospital overload with a lack of beds, we would
112 have expected a decrease in the proportion of infected individuals who are hospitalized.
113 In case of hospital overload with sufficient additional beds available but a decrease
114 in the quality of care, we would have expected IHR to remain stable with incidence
115 (since infected individuals could have been hospitalized without restriction). On the
116 contrary, we found that an incidence shift from 3% to 9% increased IHR by the same
117 magnitude as IFR (by a factor of two to three), consistent with the increase in the age
118 of infected individuals with incidence (since both IHR and IFR increase with age) [20].
119 Previous studies have compared IFR between countries and have found that age-
120 specific IFR was the main predictor of country-level IFR [11]. However, IFR varied
121 strongly between countries (by a factor of more than 30 in [11]), even after account-
122 ing for the age of the infected individuals [13, 14]. In particular, European countries
123 faced higher IFRs than expected [11–13], suggesting the importance of other determi-
124 nants of IFR than age and health-care capacity [11]. Focusing on a single country was

125 a first strength of our study, as it eliminated the influence of any country-level con-
126 founder (heterogeneity in populations, healthcare systems, mitigation policies, etc.),
127 by design. Furthermore, the association between departmental incidence and IFR was
128 not explained by the covariates we used: baseline hospital resources (measured as the
129 number of intensive care beds per inhabitant), proportion of the population over 60,
130 or prevalence of diabetes (used as a proxy for obesity). This association between inci-
131 dence and adjusted IFR was also suggested in one international comparison (see the
132 Figure 3 of [14]).

133 Some other studies found a positive association between a higher case fatality rate
134 (CFR) and hospital overload in the US and in France [4, 6]. In these studies, hospital
135 overload was measured either by the number of hospitalizations for COVID-19 or by
136 the ratio between COVID-19 cases and hospital resources, such as ICU beds or nurs-
137 ing staff. In one of these studies, CFR was standardized for the age of the population,
138 but not for the age of the infected individuals [4]. These findings may be explained
139 by the role of incidence. Indeed, hospital overload increases with incidence, as does
140 the age of infected individuals (and consequently, IFR). Incidence could therefore be
141 a confounder for hospital overload and IFR.

142 An important strength of this study was the Bayesian statistical framework, which
143 accounted for uncertainty in the latent variables used in the regressions (such as inci-
144 dence or IFR). Indeed, underestimation of incidence results in overestimation of IFR
145 (since the number of COVID-19 related deaths is known), and overestimation of inci-
146 dence leads to underestimation of IFR. Thus, sampling variation contributes to a
147 spurious negative association between incidence and IFR, as evidenced in the pos-
148 terior predictive checks (see the Supplementary information file). Considering these
149 latent variables as known during the regression step would therefore have led to biased
150 results.

151 A limitation of our study was that focusing on a single country could reduce exter-
152 nal validity, particularly for countries on other continents or with different levels of
153 development. A second limitation relates to the causal assumptions we have made,
154 as they result in a simplified view of the relation between the variables. We used the
155 prevalence of diabetes as a surrogate for obesity because it is easier to collect data on
156 (through diabetes medication), and not all risk factors for COVID-19 death were con-
157 sidered (like immunosuppression). However, these assumptions enabled formal causal
158 reasoning using graphs, which is not yet widely practiced in health research [21]. This
159 formalism contributes to greater clarity regarding the questions asked and the mech-
160 anisms under study.

161 In conclusion, this study found that a higher incidence increased the risk of death
162 among individuals infected with COVID-19. However, the mechanism of this effect
163 could be related to a shift in the profile of infected individuals with incidence. Indeed,
164 the departments with lower incidence tended to have a younger infected population (to
165 be distinguished from the age of the general population of the department). These find-
166 ings prompt a reinterpretation of studies that have observed an association between
167 the number of COVID-19 cases in a specific location and the risk of death among
168 infected individuals, as the age of infected persons was generally not considered. It is

169 important to emphasize that while we did not find epidemiological evidence implicat-
170 ing hospital overload in increasing the risk of death among COVID-19 patients, this
171 does not negate the possibility of such an effect.

172 4 Methods

173 4.1 SAPRIS-SERO study

174 This study used the data of the SAPRIS-SERO serosurvey, previously described [16,
175 22, 23]. SAPRIS-SERO was built on SAPRIS ("SAnté, Perception, pratiques, Rela-
176 tions et Inégalités Sociales en population générale pendant la crise COVID-19"), a
177 cohort whose inclusions began in March 2020, which studied epidemiological and soci-
178 ological aspects of the COVID-19 epidemic in France [22]. The adult participants of
179 SAPRIS were recruited from three cohorts based on the general population (without
180 particular selection on a disease):

- 181 • The cohort NutriNet-Santé focused on nutrition, with online follow-up. It included
182 170,000 participants at the start of the study in 2009 [24].
- 183 • The cohort CONSTANCES was set up in 2012 and included 204,973 adults,
184 selected to be a representative sample of the French adult population [25].
- 185 • E3N/E4N is a multi-generational adult cohort including 113,000 persons: the
186 women recruited at the start of the study (1990), their children, and the fathers
187 of these children [26].

188 All participants from the initial cohorts who had regular internet access and were still
189 being followed in 2020 were invited to participate in the SAPRIS study, which involved
190 self-administered questionnaires during the first wave. These questionnaires covered
191 demographic information and the history of SARS-CoV-2 testing by RT-PCR. A total
192 of 93,610 SAPRIS participants were over 20, completed the questionnaires, and resided
193 in metropolitan France. These participants were then invited to join the SAPRIS-
194 SERO study by collecting a single dried-blood spot sample themselves. The samples
195 were sent to a virology laboratory (Unité des virus émergents, Marseille, France) for
196 serological analysis using the commercial ELISA test (Euroimmun, Lübeck, Germany),
197 which detects anti-SARS-CoV-2 IgG antibodies targeting the S1 domain of the spike
198 protein. The ELISA assays performed on dried-blood spot samples demonstrated a
199 sensitivity of 98.1% to 100% and a specificity of 99.3% to 100% when compared to
200 conventional serum assays as a standard [27, 28].

201 4.2 External data

202 The results of two independent French seroprevalence studies were used as prior dis-
203 tributions for national seroprevalence (metropolitan France) [15, 29]. The results of a
204 diagnostic study concerning the ELISA test (IgG anti-S1 from Euroimmun) were used
205 as prior distributions for sensitivity and specificity [30].

206 The French population structure by age and by administrative department came from
207 the census of January first, 2020 (Insee, Institut national de la statistique et des études
208 économiques) [31]. The data about COVID-19-related hospitalizations during the first
209 semester (before July first, 2020) by administrative department were obtained from

210 the SI-VIC database, the exhaustive national inpatient surveillance system used dur-
211 ing the pandemic [32]. The data about general population mortality (including deaths
212 occurring in nursing homes) attributed to COVID-19 during the first semester (before
213 July first, 2020) were obtained from the CépiDc (Centre d'épidémiologie sur les causes
214 médicales de décès), online (open data) or directly [33]. Raw diabetes prevalence in
215 French departments in 2019 (pre-pandemic) was provided by Santé Publique France
216 (open data), based on an exhaustive monitoring of anti-diabetic drugs use (Système
217 national des données de santé) [34]. The number of ICU (intensive care unit) beds
218 per inhabitant in 2019 (pre-pandemic) was obtained from the DREES (Direction de
219 la recherche, des études, de l'évaluation et des statistiques) [35].

220 4.3 Choice of the covariates

Figure 4 features a causal graph representing departmental incidence (X), IFR or

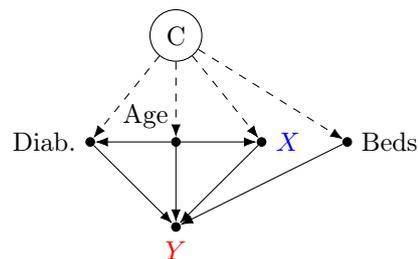


Fig. 4: Causal graph. The variables are considered at the departmental scale. The effect of X on Y can be estimated by adjusting on $\{\text{Diab.}, \text{Age}, \text{Beds}\}$.

X : COVID-19 incidence.

Y : IFR (infection fatality rate) or IHR (infection hospitalization rate).

Diab.: Prevalence of diabetes.

Beds: Number of intensive care beds per inhabitant.

Age: Proportion of population over 60.

Dashed arrows represent the effects of unmeasured confounders C .

221 IHR (Y), unobserved socio-economic variables (C), and the determinants of COVID-
222 19 outcome (according to [6, 7, 17, 18]). Our most critical assumptions, which are
223 included in the graph, are:

- 225 • Age and diabetes are the main individual risk factors for COVID-19 severity,
226 influencing hospitalization and mortality.
- 227 • The determinants of IFR and IHR at the departmental scale act through the
228 prevalence of these individual risk factors, through incidence, or through the
229 number of intensive care beds per inhabitant (the latter having a potentially
230 decisive role for IFR but acting as a surrogate for hospital beds when considering
231 IHR).

232 Given this causal graph (Figure 4), the set $\{\text{Diab.}, \text{Age}, \text{Beds}\}$ satisfies the back-
233 door criterion relative to (X, Y) and allows for estimating the causal effect of X

234 on Y despite the presence of unobserved socio-economic variables, using the back-
 235 door adjustment formula (equation 15) [36]. Prevalence of diabetes was chosen as a
 236 surrogate for obesity because it is easier to quantify precisely (through data on the
 237 sale of diabetes medication).

238 4.4 Statistical model

239 The statistical analysis was carried out within a Bayesian framework. An overview of
 240 the model is featured in Figure 5, where the equations of the model are referenced next
 to the variables. In the remainder of this section, prior distributions are not always

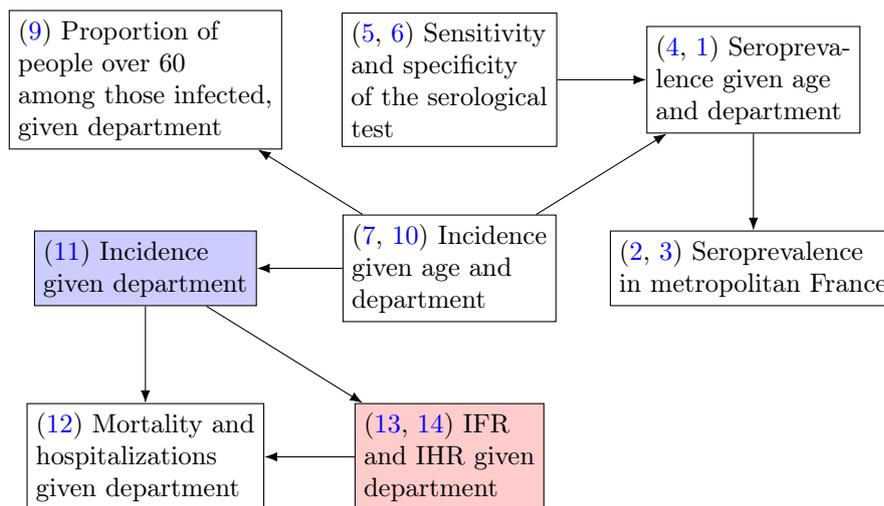


Fig. 5: Overview of the model. The blue and red rectangles represent the exposure and outcome of the main analysis, respectively. The numbers indicate the equations associated with the variables (see the Model section).

IFR: Infection fatality rate.

IHR: Infection hospitalization rate.

241 explicitly written. If so, the latter are uniform. Age groups are indexed by the letter
 242 i ($i = 0$ for individuals aged 20 to 59, and $i = 1$ for individuals aged over 60). The
 243 departments are indexed by the letter j , ranging from 1 to 95.

244 For an age group i and a department j , the participants without a positive RT-PCR
 245 nor missing data on department contributed to the estimation of seroprevalence $s_{i,j}$,
 246 considering $N_{i,j}$ the number of these participants, and $y_{i,j}$ the number of positive
 247 serological tests:
 248

$$y_{i,j} \sim B(N_{i,j}, s_{i,j}) \quad (1)$$

249 Seroprevalence at the scale of metropolitan France ($\text{sero}_{\text{France}}$) was obtained by post-
 250 stratification from $s_{i,j}$ and $\text{pop}_{i,j}$, the size of the population corresponding to this

251 group:

$$\text{sero}_{\text{France}} = \frac{\sum_{i,j} \text{pop}_{i,j} \times s_{i,j}}{\sum_{i,j} \text{pop}_{i,j}} \quad (2)$$

252 Seroprevalence estimates from other surveys were incorporated using beta distribu-
253 tions:

$$\begin{aligned} \text{sero}_{\text{France}} &\sim \text{Beta}(101, 1948) \quad (\text{implies a 95\% CI of 4.02-5.89\% [15]}) \\ \text{sero}_{\text{France}} &\sim \text{Beta}(1147, 17212) \quad (\text{implies a 95\% CI of 5.90-6.60\% [29]}) \end{aligned} \quad (3)$$

254 For an age group i and a department j , incidence (cumulated over the first semester)
255 was denoted $p_{i,j}$. Seroprevalence $s_{i,j}$ was linked to incidence $p_{i,j}$ and to the sensitivity
256 (Se) and specificity (Sp) of the serological test:

$$s_{i,j} = \text{Se} \times p_{i,j} + (1 - \text{Sp}) \times (1 - p_{i,j}) \quad (4)$$

257 Prior distributions for sensitivity and specificity originated from [30]:

$$\begin{aligned} \text{Se} &\sim \text{Beta}(585, 56) \quad (\text{implies a 95\% CI of 89.0-93.3\%}) \\ \text{Sp} &\sim \text{Beta}(953, 15) \quad (\text{implies a 95\% CI of 97.6-99.1\%}) \end{aligned} \quad (5)$$

258 The participants with a positive RT-PCR contributed to the likelihood of sensitivity.
259 With N_{se} and y_{se} the number of total and positive (respectively) serological tests in
260 this group,

$$y_{\text{se}} \sim \text{B}(N_{\text{se}}, \text{Se}) \quad (6)$$

261 Incidence $p_{i,j}$ was modeled on the logit scale. As it was suggested that older persons
262 could be under-represented in the population of infected persons at the early stages
263 of the epidemic [19], every department had a unique log-odds ratio for age over 60
264 (β_j), possibly influenced by its intercept α_j (logit of incidence in the 20-59 in the
265 department j) through a linear regression with intercept μ_{age} , slope b_{age} and standard
266 deviation σ_{age} (hierarchical modeling, β_j is a random effect):

$$\begin{aligned} \text{logit } p_{i,j} &= \alpha_j + \underbrace{\beta_j}_{\text{random slope}} \times i \\ \beta_i &\sim N(\mu_{\text{age}} + \alpha_j \times b_{\text{age}}, \sigma_{\text{age}}) \end{aligned} \quad (7)$$

267 The α_j departmental intercepts entailed spatial auto-correlation through an ICAR
268 (intrinsic conditional auto-regressive) component ϕ_j (as described and implemented in
269 this reference [37]), associated with an overall intercept μ_{α} and a scale parameter σ_{ϕ}
270 representing the amount of spatial correlation:

$$\begin{aligned} \alpha_j &= \mu_{\alpha} + \phi_j \times \sigma_{\phi} \\ \sigma_{\phi} &\sim \text{Exponential}(1) \quad (\text{weakly informative prior on } \sigma_{\phi}) \end{aligned} \quad (8)$$

271 The proportion $\text{age}_{\text{infected},j}$ of persons over 60 among those infected in a department
272 j was reconstructed from $p_{i,j}$ and from $\text{age}_{\text{pop},j} = \frac{\text{pop}_{1,j}}{\text{pop}_{1,j} + \text{pop}_{0,j}}$ (the proportion of
273 persons above 60 in the population of the department j):

$$\text{age}_{\text{infected},j} = \frac{p_{1,j} \times \text{age}_{\text{pop},j}}{p_{1,j} \times \text{age}_{\text{pop},j} + p_{0,j} \times (1 - \text{age}_{\text{pop},j})} \quad (9)$$

274 For a given incidence in the persons under 60, we estimated an expected proportion
275 of persons over 60 among those infected for a department with the same age structure
276 as metropolitan France. This expected proportion was reconstructed from the coef-
277 ficients μ_{age} , β_{age} , and σ_{age} (equation 7), according to a procedure described in the
278 Supplementary information file ("Computation of expectations" section). This anal-
279 ysis aimed to illustrate the dependence between incidence in people under 60 and
280 incidence in those over 60, and its possible consequences for IFR.

281 The K participants with a positive RT-PCR and no missing data concerning the
282 department contributed directly to incidence. With IS_k being the infection status of
283 the participant k (IS_k is always equal to 1 in this positive RT-PCR group),

$$\text{IS}_k \sim \text{Bern}(p_{i,j}) \quad (\text{For } k = 1, \dots, K) \quad (10)$$

284 Departmental incidence inc_j was obtained by post-stratification from $p_{i,j}$ and $\text{pop}_{i,j}$:

$$\text{inc}_j = \frac{\sum_i \text{pop}_{i,j} \times p_{i,j}}{\sum_i \text{pop}_j} \quad (11)$$

285 For a department j , the counts of deaths (D_j) and hospitalizations (H_j) of the first
286 semester were modeled with Poisson regressions:

$$\begin{aligned} D_j &\sim \text{Poisson}(\text{IFR}_j \times \text{inc}_j \times \underbrace{\sum_i \text{pop}_{i,j}}_{\text{n. infected}}) \\ H_j &\sim \text{Poisson}(\text{IHR}_j \times \text{inc}_j \times \underbrace{\sum_i \text{pop}_{i,j}}_{\text{n. infected}}) \end{aligned} \quad (12)$$

287 Departmental IFRs and IHRs were modeled as logistic functions of linear predictors
288 $g_{\text{IFR}}(j)$ and $g_{\text{IHR}}(j)$, ranging respectively from 0% to 5% and from 0% to 10%:

$$\begin{aligned} \text{IFR}_j &= \frac{0.05}{1 + e^{-g_{\text{IFR}}(j)}} \quad (\text{logistic function ranging from 0\% to 5\%}) \\ \text{IHR}_j &= \frac{0.10}{1 + e^{-g_{\text{IHR}}(j)}} \quad (\text{logistic function ranging from 0\% to 10\%}) \end{aligned} \quad (13)$$

289 These linear predictors included a departmental random intercept, a slope represent-
290 ing the role of incidence, and coefficients associated with the covariates (prevalence of

291 diabetes, number of intensive care beds per inhabitant and proportion of the popula-
292 tion over 60). With $diab_j$, $age_{pop,j}$, and $beds_j$ being the covariates for the department
293 j ,

$$\begin{aligned}
 g_{IFR}(j) &= \overbrace{\mu_{IFR_j}}^{\text{random intercept}} + \underbrace{inc_j \times c_p}_{\text{role of incidence}} + \overbrace{diab_j \times c_{diab} + age_{pop,j} \times c_{age} + beds_j \times c_{beds}}^{\text{role of the covariates}} \\
 g_{IHR}(j) &= \overbrace{\mu_{IHR_j}}^{\text{random intercept}} + \underbrace{inc_j \times d_p}_{\text{role of incidence}} + \overbrace{diab_j \times d_{diab} + age_{pop,j} \times d_{age} + beds_j \times d_{beds}}^{\text{role of the covariates}} \quad (14) \\
 \mu_{IFR_j} &\sim N(m_{IFR}, \sigma_{IFR}) \\
 \mu_{IHR_j} &\sim N(m_{IHR}, \sigma_{IHR}) \\
 \sigma_{IFR} &\sim \text{Exponential}(1) \text{ (weakly informative prior on } \sigma_{IFR}\text{)} \\
 \sigma_{IHR} &\sim \text{Exponential}(1) \text{ (weakly informative prior on } \sigma_{IHR}\text{)}
 \end{aligned}$$

294 Because the covariates satisfy the back-door criterion, the adjusted IFR can be equated
295 to the causal effect of incidence on IFR using the back-door adjustment formula [36].
296 $P(IFR|do(inc = x))$ denotes this causal effect, which is the distribution of departmen-
297 tal IFRs if incidence was artificially set to the value x without any other modification.
298 With z_j denoting the vector of covariates for department j , and because the distri-
299 bution of these covariates in the French departments is known ($P(z_j) = \frac{1}{95}$ for all
300 j),

$$\begin{aligned}
 \underbrace{P(IFR|do(inc = x))}_{\text{causal effect of incidence } x \text{ on IFR}} &= \overbrace{\sum_{j=1}^{95} P(IFR|x, z_j) \times P(z_j)}^{\text{adjusted IFR given incidence } x} \\
 &= \frac{1}{95} \sum_{j=1}^{95} P(IFR|x, z_j) \quad (15)
 \end{aligned}$$

301 The coefficients of equation 14 were used to compute $E[IFR|do(inc = x)]$, as described
302 in the supplementary information file ("Reconstruction of expectations" section). The
303 target of this study was the average causal effect of an incidence shift from 3% to 9%:

$$E[IFR|do(inc = 9\%)] - E[IFR|do(inc = 3\%)] \quad (16)$$

304 This average causal effect corresponds to the expected difference in IFR when arti-
305 ficially setting incidence to 9% versus 3% in a department of metropolitan France
306 (without changing anything else than incidence). The same procedure was applied to
307 estimate the average causal effect of this incidence shift on IHR.

308 4.5 Algorithm and software

309 The data management was done using R version 4.3.1, and the modeling was per-
310 formed with Stan (R package cmdstanr version 0.5.3), which implements Hamiltonian
311 Monte Carlo (HMC) [38, 39]. The models for the random coefficients (β_j , μ_{IFR_j} and

312 μ_{IFR_j}) employed non-centered parameterizations to improve HMC convergence [40].
313 The Monte Carlo sampling consisted of 8 chains of 2,000 iterations each (includ-
314 ing 1,000 warm up iterations). Trace plots, \hat{R} statistics and effective Monte Carlo
315 sample sizes provided by Stan were used to assess convergence. Posterior predic-
316 tive checks are provided in the Supplementary information file. The model's code (in
317 Stan) is provided in the Supplementary information file and in a GitHub repository
318 (<https://github.com/bglemain/does-hospital-overload>).

319 4.6 Ethical approval and consent to participate

320 Ethical approval and written or electronic informed consent were obtained from each
321 participant before enrollment in the original cohort. The SAPRIS-SERO study was
322 approved by the Sud-Méditerranée III ethics committee (approval 20.04.22.74247) and
323 electronic informed consent was obtained from all participants for dried blood spot
324 testing. The study was registered (#NCT04392388). All methods were performed in
325 accordance with the relevant guidelines and regulations.

326 **Supplementary information.** The Supplementary information file attached with
327 this article contains:

- 328 • A flowchart
- 329 • The number of serological samples for each department and age group
- 330 • The exhaustive list of departmental estimates (incidence, IFR, IHR) and charac-
331 teristics (covariates, population size, counts of deaths and hospitalizations)
- 332 • Complementary results (univariate analysis, role of the confounders)
- 333 • The procedure used to compute expectations after a logistic transformation
- 334 • Posterior predictive checks
- 335 • The Stan code

336 **Acknowledgments.** The authors warmly thank all the volunteers of the Con-
337 stances, E3N–E4N, and NutriNet-Santé cohorts. We thank the staff of the Constances,
338 E3N–E4N and NutriNet-Santé cohorts that have worked with dedication and engage-
339 ment to collect and manage the data used for this study and to ensure continuing
340 communication with the cohort participants. We thank the CEPH-Biobank staff for
341 their adaptability and the quality of their work. We thank all the members of the
342 SAPRIS-SERO study group (see Supplementary information file).

343 Declarations

- 344 • Funding: This study ANR (Agence Nationale de la Recherche, #ANR-20-COVI-
345 000, #ANR-10-COHO-06), Fondation pour la Recherche Médicale (#20RR052-
346 00), Inserm (Institut National de la Santé et de la Recherche Médicale, #C20-26).
347 The sponsor and funders facilitated data acquisition but did not participate
348 in the study design, analysis, interpretation or drafting. Cohorts funding The
349 CONSTANCES Cohort Study is supported by the Caisse Nationale d'Assurance
350 Maladie (CNAM), the French Ministry of Health, the Ministry of Research, the
351 Institut national de la santé et de la recherche médicale. CONSTANCES benefits

352 from a grant from the French National Research Agency [Grant Number ANR-
353 11-INBS-0002] and is also partly funded by MSD, AstraZeneca, Lundbeck and
354 L’Oreal. The E3N-E4N cohort is supported by the following institutions: Min-
355 istère de l’Enseignement Supérieur, de la Recherche et de l’Innovation, INSERM,
356 University Paris-Saclay, Gustave Roussy, the MGEN, and the French League
357 Against Cancer. The NutriNet-Santé study is supported by the following public
358 institutions: Ministère de la Santé, Santé Publique France, Institut National de la
359 Santé et de la Recherche Médicale (INSERM), Institut National de la Recherche
360 Agronomique (INRAE), Conservatoire National des Arts et Métiers (CNAM)
361 and Sorbonne Paris Nord. The CEPH-Biobank is supported by the Ministère de
362 l’Enseignement Supérieur, de la Recherche et de l’Innovation.

- 363 • Competing interests: The authors declare no competing interests.
- 364 • Ethics approval: Ethical approval and written or electronic informed consent
365 were obtained from each participant before enrolment in the original cohort. The
366 SAPRIS-SERO study was approved by the Sud-Méditerranée III ethics committee
367 (approval #20.04.22.74247)
- 368 • Consent to participate: An electronic informed consent was obtained from all
369 participants for DBS testing
- 370 • Consent for publication: Participants can not be identified on the basis of this
371 article
- 372 • Availability of data and materials: In regard to data availability, data from the
373 study are protected under the protection of health data regulation set by the
374 French National Commission on Informatics and Liberty (Commission Nationale
375 de l’Informatique et des Libertés, CNIL). The data can be made available upon
376 reasonable request to F.C. (fabrice.carrat@iplesp.upmc.fr), after a consultation
377 with the steering committee of the SAPRIS-SERO study. The French law forbids
378 us to provide free access to SAPRIS-SERO data; access could however be given
379 by the steering committee after legal verification of the use of the data. Please,
380 feel free to come back to us should you have any additional question.
- 381 • Code availability: The code Stan for the model is provided in Supplementary
382 information file
- 383 • Authors’ contributions: F.C., N.L., C.A., W.G., P.M, and B.G. conceived and
384 designed the study. B.G. implemented the model and wrote the manuscript. All
385 the authors reviewed and edited the manuscript.

386 Consortium (SAPRIS-SERO study group)

387 Fabrice Carrat^{1,2}, Pierre-Yves Ancel¹¹, Marie-Aline Charles¹¹, Gianluca Severi^{7,8},
388 Mathilde Touvier⁹, Marie Zins^{5,6}, Sofiane Kab⁶, Adeline Renvy⁶, Stephane Le-Got⁶,
389 Celine Ribet⁶, Mireille Pellicer⁶, Emmanuel Wiernik⁶, Marcel Goldberg⁶, Fanny
390 Artaud⁷, Pascale Gerbouin-Rérolle⁷, Mélody Enguix⁷, Camille Laplanche⁷, Rose-
391 lyn Gomes-Rima⁷, Lyan Hoang⁷, Emmanuelle Correia⁷, Alpha Amadou Barry⁷,
392 Nadège Senina⁷, Julien Allegre⁹, Fabien Szabo de Edelenyi⁹, Nathalie Druesne-
393 Pecollo⁹, Younes Esseddik⁹, Serge Herberg⁹, Mélanie Deschasaux⁹, Marie-Aline
394 Charles¹¹, Valérie Benhammou¹², Anass Ritmi¹³, Laetitia Marchand¹³, Cecile Zaros¹³,

395 Elodie Lordmi¹³, Adriana Candea¹³, Sophie de Visme¹³, Thierry Simeon¹³, Xavier
396 Thierry¹³, Bertrand Geay¹³, Marie-Noelle Dufourg¹³, Karen Milcent¹³, Delphine
397 Rahib¹⁴, Nathalie Lydie¹⁴, Clovis Lusivika-Nzinga¹, Gregory Pannetier¹, Nathanael
398 Lapidus^{1,2}, Isabelle Goderel¹, Céline Dorival¹, Jérôme Nicol¹, Olivier Robineau¹,
399 Cindy Lai¹⁵, Liza Belhadji¹⁵, Hélène Esperou¹⁵, Sandrine Couffin-Cadiergues¹⁵,
400 Jean-Marie Gagliolo¹⁶, Hélène Blanché¹⁰, Jean-Marc Sébaoun¹⁰, Jean-Christophe
401 Beauoin¹⁰, Laetitia Gressin¹⁰, Valérie Morel¹⁰, Ouissam Ouili¹⁰, Jean-François
402 Deleuze¹⁰, Laetitia Ninove⁴, Stéphane Priet⁴, Paola Mariela Saba Villarroel⁴, Toscane
403 Fourié⁴, Souand Mohamed Ali⁴, Abdenour Amroun⁴, Morgan Seston⁴, Nazli Ayhan⁴,
404 Boris Pastorino⁴, Xavier de Lamballerie⁴.

405 ¹Sorbonne Université, Inserm, Institut Pierre-Louis d'épidémiologie et de santé
406 publique, Paris, France

407 ²Département de santé publique, Hôpital Saint-Antoine, AP-HP.Sorbonne université,
408 Paris, France

409 ⁴Unité des Virus Émergents, UVE, Aix Marseille Univ, IRD 190, INSERM 1207, IHU
410 Méditerranée Infection, Marseille, France

411 ⁵Paris University, Paris, France

412 ⁶Université Paris-Saclay, Université de Paris, UVSQ, Inserm UMS 11, Villejuif, France

413 ⁷CESP UMR1018, Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, Villejuif,
414 France

415 ⁸Department of Statistics, Computer Science and Applications, University of Florence,
416 Italy

417 ⁹Sorbonne Paris Nord University, Inserm U1153, Inrae U1125, Cnam, Nutritional Epi-
418 demiology Research Team (EREN), Epidemiology and Statistics Research Center –
419 University of Paris (CRESS), Bobigny, France

420 ¹⁰Fondation Jean Dausset-CEPH (Centre d'Etude du Polymorphisme Humain),
421 CEPH-Biobank, Paris, France

422 ¹¹Centre for Research in Epidemiology and StatisticS (CRESS), Inserm, INRAE, Uni-
423 versité de Paris, Paris, France

424 ¹²EPIPAGE-2 Joint Unit, Paris, France

425 ¹³ELFE Joint Unit, Paris, France

426 ¹⁴Santé Publique France, Paris, France

427 ¹⁵Inserm, Paris, France

428 ¹⁶Aviesan, Inserm, Paris, France

429 References

430 [1] Candel, F. J. et al. Temporary hospitals in times of the COVID pandemic. An
431 example and a practical view. *Rev. Esp. Quimioter.* **34**, 280 (2021).

432 [2] Lefrant, J.-Y. et al. A national healthcare response to intensive care bed require-
433 ments during the COVID-19 outbreak in France. *Anaesth Crit Care Pain Med*
434 **39**, 709–715 (2020).

- 435 [3] Jimenez, J. V. et al. Outcomes in Temporary ICUs Versus Conventional ICUs:
436 An Observational Cohort of Mechanically Ventilated Patients With COVID-19–
437 Induced Acute Respiratory Distress Syndrome. *Crit Care Explor* **4**, e0668 (2022).
- 438 [4] Souris, M. & Gonzalez, J.-P. COVID-19: Spatial analysis of hospital case-fatality
439 rate in France. *PLoS One* **15**, e0243606 (2020).
- 440 [5] Zappella, N. et al. Temporary ICUs during the COVID-19 pandemic first wave:
441 Description of the cohort at a French centre. *BMC Anesthesiol* **22**, 310 (2022).
- 442 [6] Janke, A. T. et al. Analysis of Hospital Resource Availability and COVID-19
443 Mortality Across the United States. *J Hosp Med* **16**, 211–214 (2021).
- 444 [7] Sen-Crowe, B., Sutherland, M., McKenney, M. & Elkbuli, A. A Closer Look Into
445 Global Hospital Beds Capacity and Resource Shortages During the COVID-19
446 Pandemic. *J Surg Res* **260**, 56–63 (2021).
- 447 [8] Dudel, C. et al. Monitoring trends and differences in COVID-19 case-fatality rates
448 using decomposition methods: Contributions of age structure and age-specific
449 fatality. *PLoS One* **15**, e0238904 (2020).
- 450 [9] Lau, H. et al. Evaluating the massive underreporting and undertesting of COVID-
451 19 cases in multiple global epicenters. *Pulmonology* **27**, 110–115 (2021).
- 452 [10] Borgdorff, M. W. New Measurable Indicator for Tuberculosis Case Detection.
453 *Emerg. Infect. Dis.* **10**, 1523 (2004).
- 454 [11] COVID-19 Forecasting Team. Variation in the COVID-19 infection-fatality ratio
455 by age, time, and geography during the pre-vaccine era: A systematic analysis.
456 *Lancet* **399**, 1469–1488 (2022).
- 457 [12] O’Driscoll, M. et al. Age-specific mortality and immunity patterns of SARS-CoV-
458 2. *Nature* **590**, 140–145 (2021).
- 459 [13] Pezzullo, A. M., Axfors, C., Contopoulos-Ioannidis, D. G., Apostolatos, A. &
460 Ioannidis, J. P. A. Age-stratified infection fatality rate of COVID-19 in the non-
461 elderly population. *Environ Res* **216**, 114655 (2023).
- 462 [14] Ioannidis, J. P. A. Infection fatality rate of COVID-19 inferred from seropreva-
463 lence data. *Bull World Health Organ* **99**, 19–33F (2021).
- 464 [15] Le Vu, S. et al. Prevalence of SARS-CoV-2 antibodies in France: Results from
465 nationwide serological surveillance. *Nat Commun* **12**, 3025 (2021).
- 466 [16] Carrat, F. et al. Antibody status and cumulative incidence of SARS-CoV-2 infec-
467 tion among adults in three regions of France following the first lockdown and
468 associated risk factors: A multicohort study. *Int J Epidemiol* **50**, 1458–1472
469 (2021).

- 470 [17] COVID-ICU Group on behalf of the REVA Network and the COVID-ICU Inves-
471 tigators. Clinical characteristics and day-90 outcomes of 4244 critically ill adults
472 with COVID-19: A prospective cohort study. *Intensive Care Med* **47**, 60–73
473 (2021).
- 474 [18] Brown, P. A. Country-level predictors of COVID-19 mortality. *Sci Rep* **13**, 9263
475 (2023).
- 476 [19] Tran Kiem, C. et al. SARS-CoV-2 transmission across age groups in France and
477 implications for control. *Nat Commun* **12**, 6895 (2021).
- 478 [20] Glemain, B. et al. Estimating SARS-CoV-2 infection probabilities with serological
479 data and a Bayesian mixture model. *Sci Rep* **14**, 9503 (2024).
- 480 [21] Tennant, P. W. G. et al. Use of directed acyclic graphs (DAGs) to identify
481 confounders in applied health research: Review and recommendations. *Int J*
482 *Epidemiol* **50**, 620–632 (2021).
- 483 [22] Carrat, F. et al. Incidence and risk factors of COVID-19-like symptoms in the
484 French general population during the lockdown period: A multi-cohort study.
485 *BMC Infect Dis* **21**, 169 (2021).
- 486 [23] Carrat, F. et al. Age, COVID-19-like symptoms and SARS-CoV-2 seropositivity
487 profiles after the first wave of the pandemic in France. *Infection* **50**, 257–262
488 (2022).
- 489 [24] Hercberg, S. et al. The Nutrinet-Santé Study: A web-based prospective study
490 on the relationship between nutrition and health and determinants of dietary
491 patterns and nutritional status. *BMC Public Health* **10**, 242 (2010).
- 492 [25] Zins, M. & Goldberg, M. The French CONSTANCES population-based cohort:
493 Design, inclusion and follow-up. *Eur J Epidemiol* **30**, 1317–1328 (2015).
- 494 [26] Clavel-Chapelon, F. & E3N Study Group. Cohort Profile: The French E3N Cohort
495 Study. *Int J Epidemiol* **44**, 801–809 (2015).
- 496 [27] Morley, G. L. et al. Sensitive Detection of SARS-CoV-2-Specific Antibodies in
497 Dried Blood Spot Samples. *Emerg Infect Dis* **26**, 2970–2973 (2020).
- 498 [28] Zava, T. T. & Zava, D. T. Validation of dried blood spot sample modifications to
499 two commercially available COVID-19 IgG antibody immunoassays. *Bioanalysis*
500 **13**, 13–28 (2021).
- 501 [29] Warszawski, J. et al. Trends in social exposure to SARS-Cov-2 in France. Evidence
502 from the national socio-epidemiological cohort–EPICOV. *PLoS One* **17**, e0267725
503 (2022).

- 504 [30] Otter, A. D. et al. Implementation and Extended Evaluation of the Euroim-
505 mun Anti-SARS-CoV-2 IgG Assay and Its Contribution to the United Kingdom's
506 COVID-19 Public Health Response. *Microbiol Spectr* **10**, e0228921 (2022).
- 507 [31] Populations légales 2020 Recensement de la population
508 Régions, départements, arrondissements, cantons et communes.
509 <https://www.insee.fr/fr/statistiques/6683031?sommaire=6683037>.
- 510 [32] Données hospitalières relatives à l'épidémie de COVID-19 (SIVIC).
511 [https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-](https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/)
512 [lepidemie-de-covid-19/](https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/).
- 513 [33] Covid-19 - Inserm-CépiDc. <https://opendata.idf.inserm.fr/cepidc/covid-19/>.
- 514 [34] Géodes - Santé publique France. <https://geodes.santepubliquefrance.fr/#c=home>.
- 515 [35] La Statistique annuelle des établissements (SAE) | Direction de la recherche,
516 des études, de l'évaluation et des statistiques. [https://drees.solidarites-](https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/00-la-statistique-annuelle-des-etablisements-sae)
517 [sante.gouv.fr/sources-outils-et-enquetes/00-la-statistique-annuelle-des-](https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/00-la-statistique-annuelle-des-etablisements-sae)
518 [etablisements-sae](https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/00-la-statistique-annuelle-des-etablisements-sae).
- 519 [36] Pearl, J. Causal Diagrams for Empirical Research. *Biometrika* **82**, 669–688
520 (1995).
- 521 [37] Morris, M. et al. Bayesian hierarchical spatial models: Implementing the Besag
522 York Mollié model in stan. *Spat Spatiotemporal Epidemiol* **31**, 100301 (2019).
- 523 [38] R Core Team. *R: A Language and Environment for Statistical Computing*. R
524 Foundation for Statistical Computing, Vienna, Austria (2023). URL [https://](https://www.R-project.org/)
525 www.R-project.org/.
- 526 [39] Carpenter, B. et al. Stan: A Probabilistic Programming Language. *J Stat Softw*
527 **76**, 1 (2017).
- 528 [40] Papaspiliopoulos, O. & Roberts, G. Non-Centered Parameterisations for Hier-
529 archical Models and Data Augmentation. *Bayesian Statistics* **7**, 307–326
530 (2003).